

Moving Science Data: One CDN to rule them all

Derek Weitzel

University of Nebraska - Lincoln

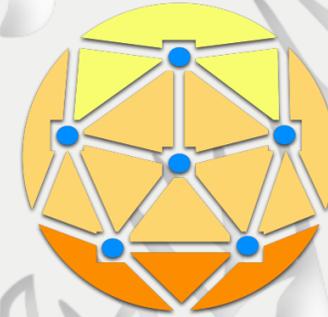
This project is supported by National Science Foundation under Cooperative Agreement OAC-2030508. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

StashCache

- The first discussion of StashCache was in January 2015

StashCache

Data delivery network
for OSG Connect users



Anna Olson • University of Chicago

XRootD Workshop
29 January 2015



Open Science Grid



StashCache – Modest Beginnings

3 Caches, 1 Origin

StashCache locations & compute sites

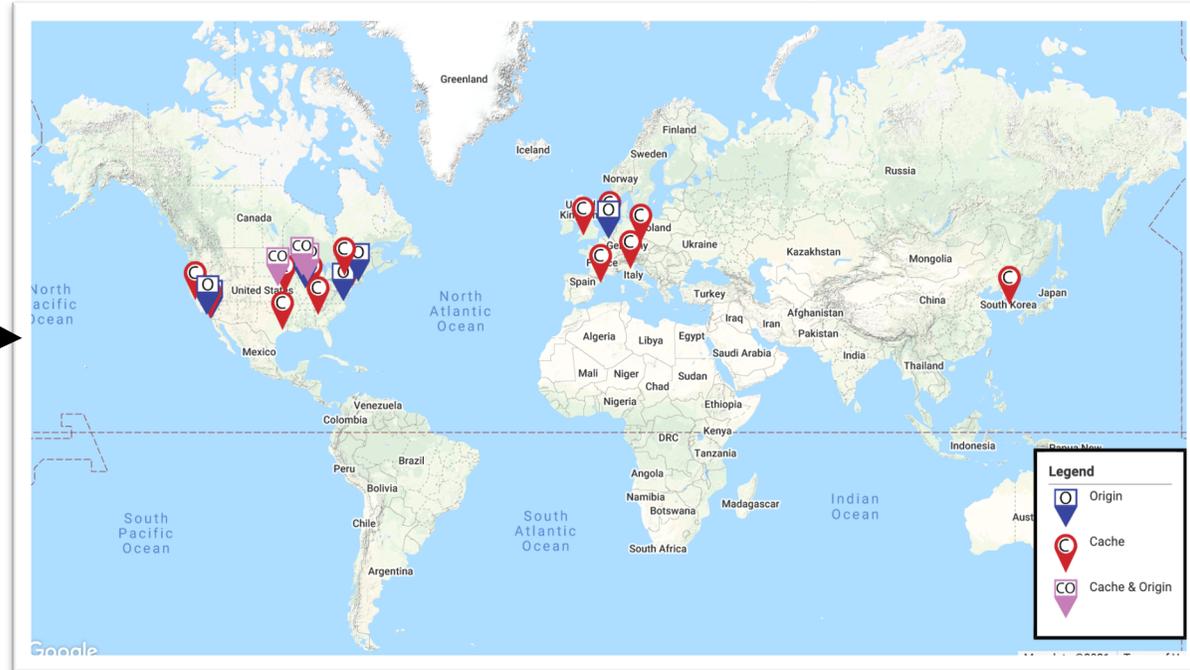
Stash
origin: ★

OSG
Caches: ●



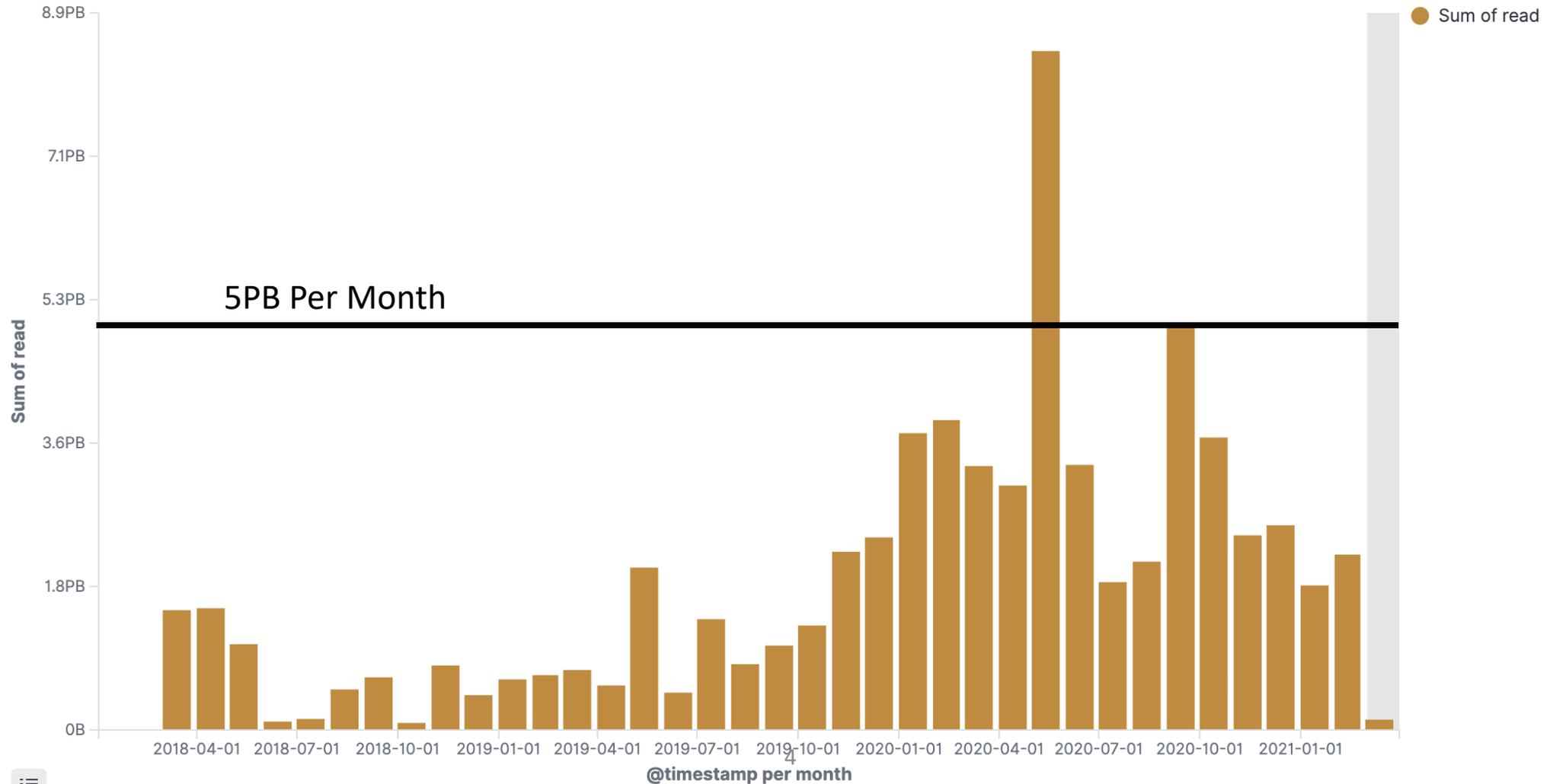
2015

17 Caches, 10 Origins



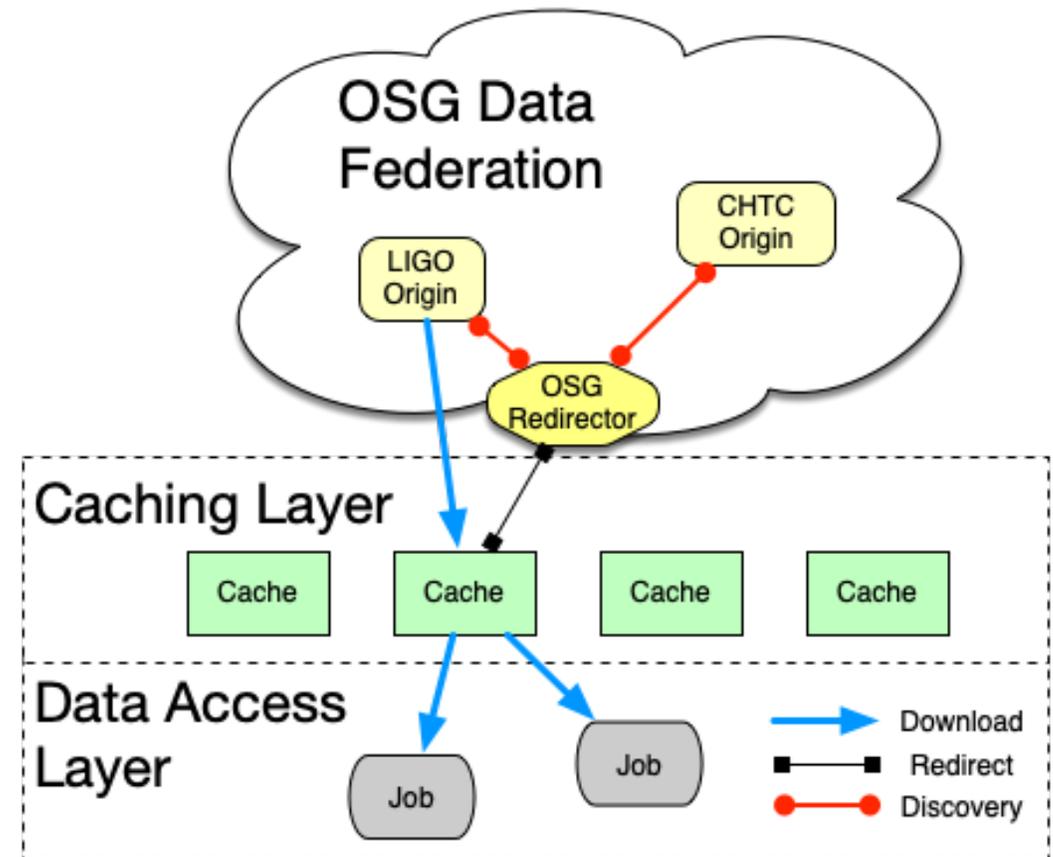
<https://map.opensciencegrid.org/>

StashCache – Growth per Month



StashCache – Overview

- Origins are operated by collaborations, or a general use OSG origin
- Origins provide data to the caching layer
- Caches are run by organizations or the OSG in strategic regional locations
- Regional caches provide data to jobs.



StashCache – Access

- Adding files to StashCache depends on the Origin and the collaboration:
- **OSG Connect:** users place files in their /public directory and it is available to the data federation.
- **LIGO:** uses a production service to copy data to the LIGO origin in Nebraska. No user generated data, only instrument data.
- **FNAL Origins:** place files in a special dCache directory.

StashCache – Access

- StashCache data is accessed primarily through 2 methods, CVMFS and StashCP
- **CVMFS:**
 - Provides a POSIX-like interface to data

```
$ cat /cvmfs/stash.osgstorage.org/osgconnect/public/dweitzel/blast/queries/query1  
>Derek's first query!  
MPVSDSGFDNSSKTMKDDTIPTEDYEEITKESEMGDATKITSKIDANVIEKKDTSENNITIAQDDEKVSWLQRVVEFFE
```

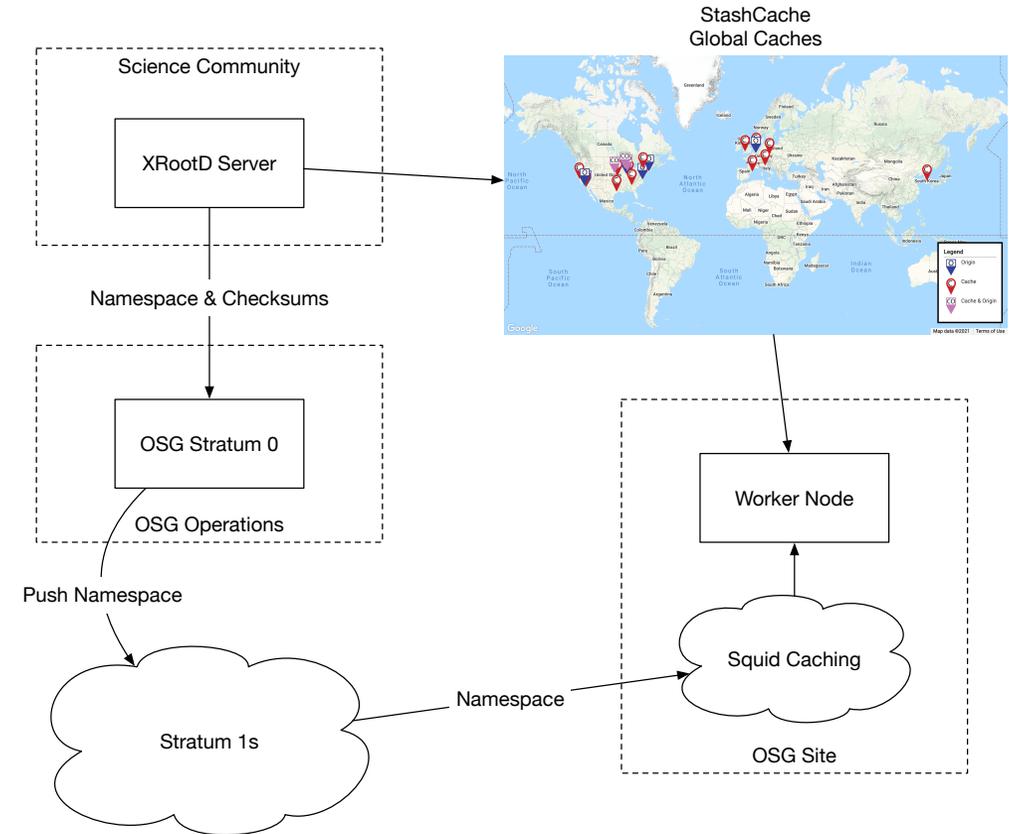
StashCache – CVMFS

```
$ cat /cvmfs/stash.osgstorage.org/osgconnect/public/dweitzel/blast/queries/query1  
>Derek's first query!  
MPVSDSGFDNSSKTMKDDTIPTEDYEEITKESEMGDATKITSKIDANVIEKKDTSENNITIAQDDEKVSWLQRVVEFFE
```

- **stash.osgstorage.org** – CVMFS Repo for users of OSG Connect
- **osgconnect** – Namespace identifier in the StashCache federation
- Path is the same as on the OSG Connect login nodes

StashCache – CVMFS

- CVMFS stores the namespace separate from the data
- Namespace is through cached HTTP
- Data is through StashCache Federation
- CVMFS can take ~8 hours to scan and update namespace with changes.



StashCache – Access

- **StashCP**
- Custom tool developed for StashCache
- Data is immediately available, no need to wait for CVMFS scan
- Tries multiple methods to copy data: HTTP, XRootD, and CVMFS (if available)
- CVMFS is not required, therefore more available resources

```
$ stashcp /osgconnect/public/dweitzel/blast/queries/query1 ./
```

StashCache – Cache Selection

- Both **CVMFS** and **StashCP** use GeolP to determine nearest cache to use.
- Both clients will fall back to other near caches if the selected cache does not respond or is too slow*

Manhattan Cache Feb. 2021

Client Domain	Bytes Read
uconn.edu	37.8TB
amnh.org	6.7TB
syr.edu	1.9TB
org.br	372.1GB
verizon.net	36.9GB
mit.edu	5.3GB
pic.es	5.2GB
rutgers.edu	3.5GB
ac.uk	1.5GB
in2p3.fr	898.2MB

Disclaimer: Client domain is difficult to capture and will not capture clients from all clusters

StashCache – Deploying

- Both Origins and Caches can be deployed using RPMs or Containers
- Origins can support multiple filesystems. In the OSG we have:
 - Regular Disk Server
 - NFS Mount
 - HDFS
 - dCache

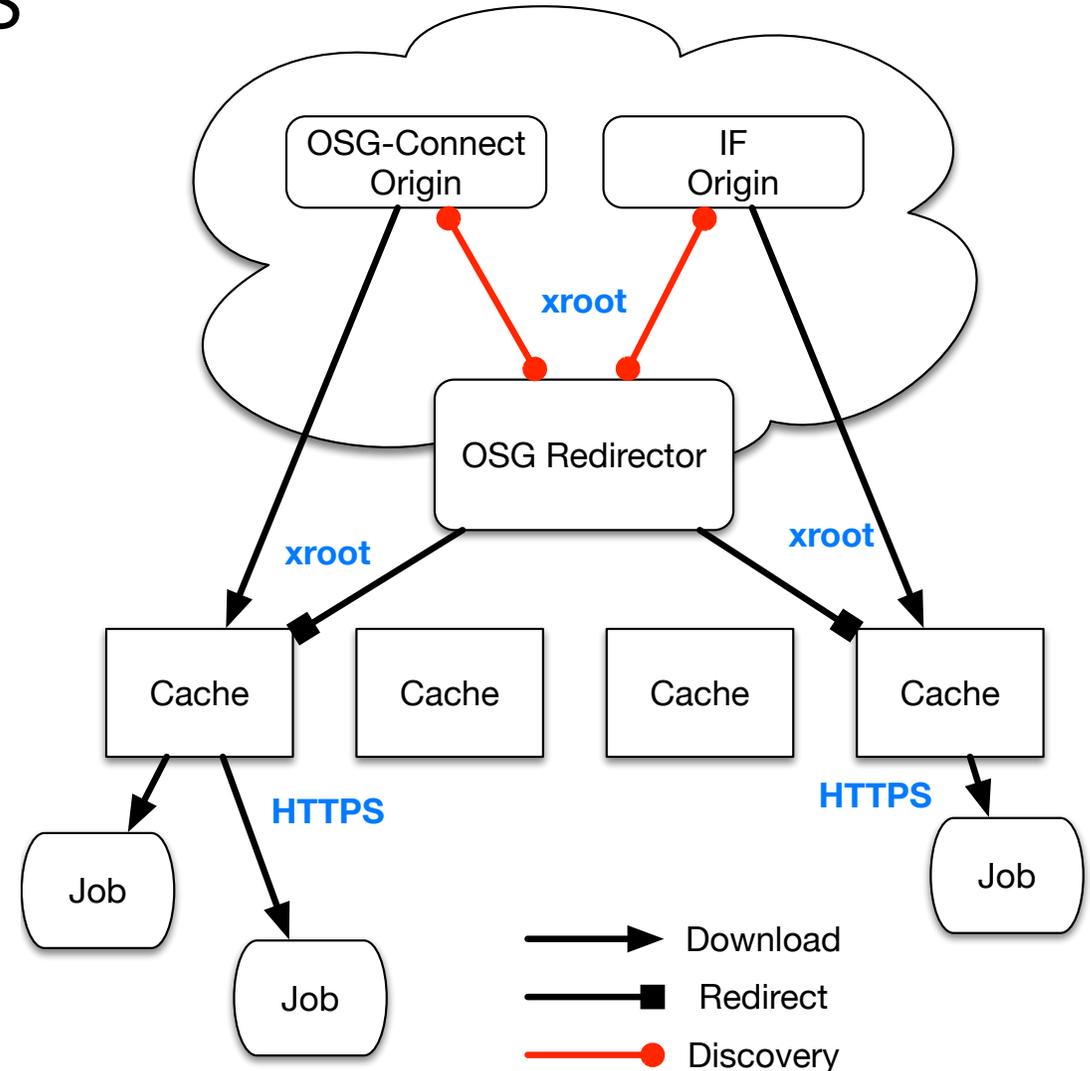
<https://opensciencegrid.org/docs/data/stashcache/overview/>

StashCache – Caches

- 5 Internet2 hosted caches – Kubernetes
- 7 U.S. OSG Contributor hosted – Mix of Kubernetes and RPM
- 5 Caches in Europe – Kubernetes

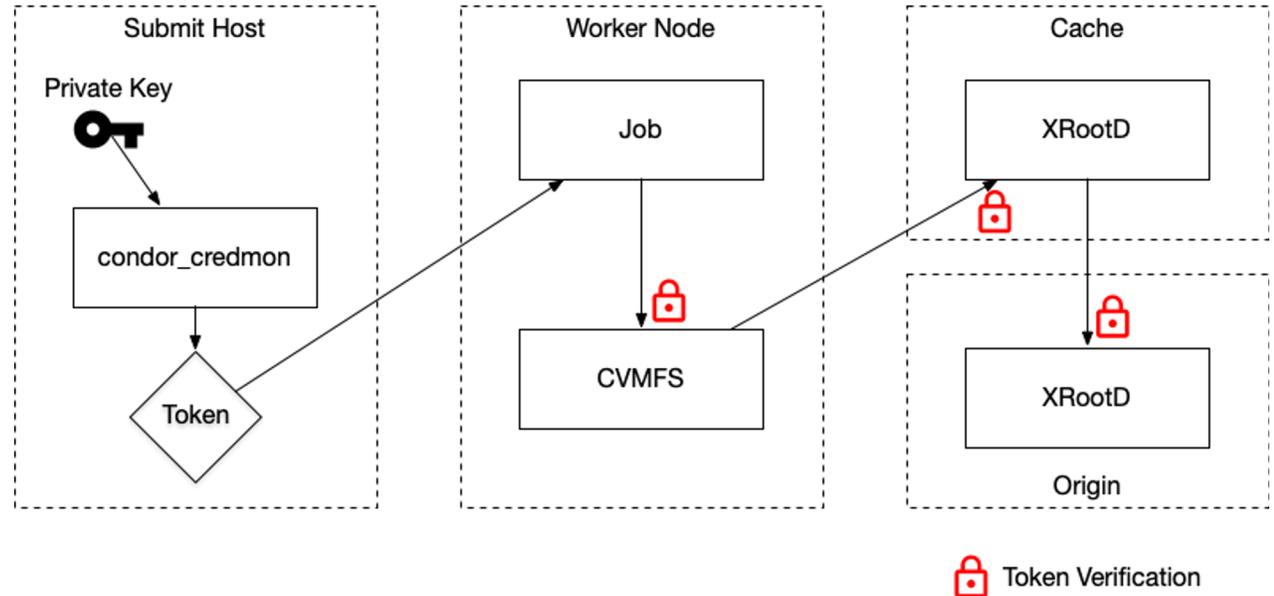
StashCache - Tokens

- The entire StashCache infrastructure is SciTokens compatible
- Must encrypt communication to secure the token
- Job to Cache is over **HTTPS**, encrypted
- **xroot** protocol between cache, redirector, and origins
- Only recently added fully encrypted communication



StashCache – Tokens

1. Token is transferred securely with job
2. CVMFS verifies token on the worker node (cache access)
3. Token is used to request data from cache
4. Cache propagates token to gather data from Origin



Accounting Validations

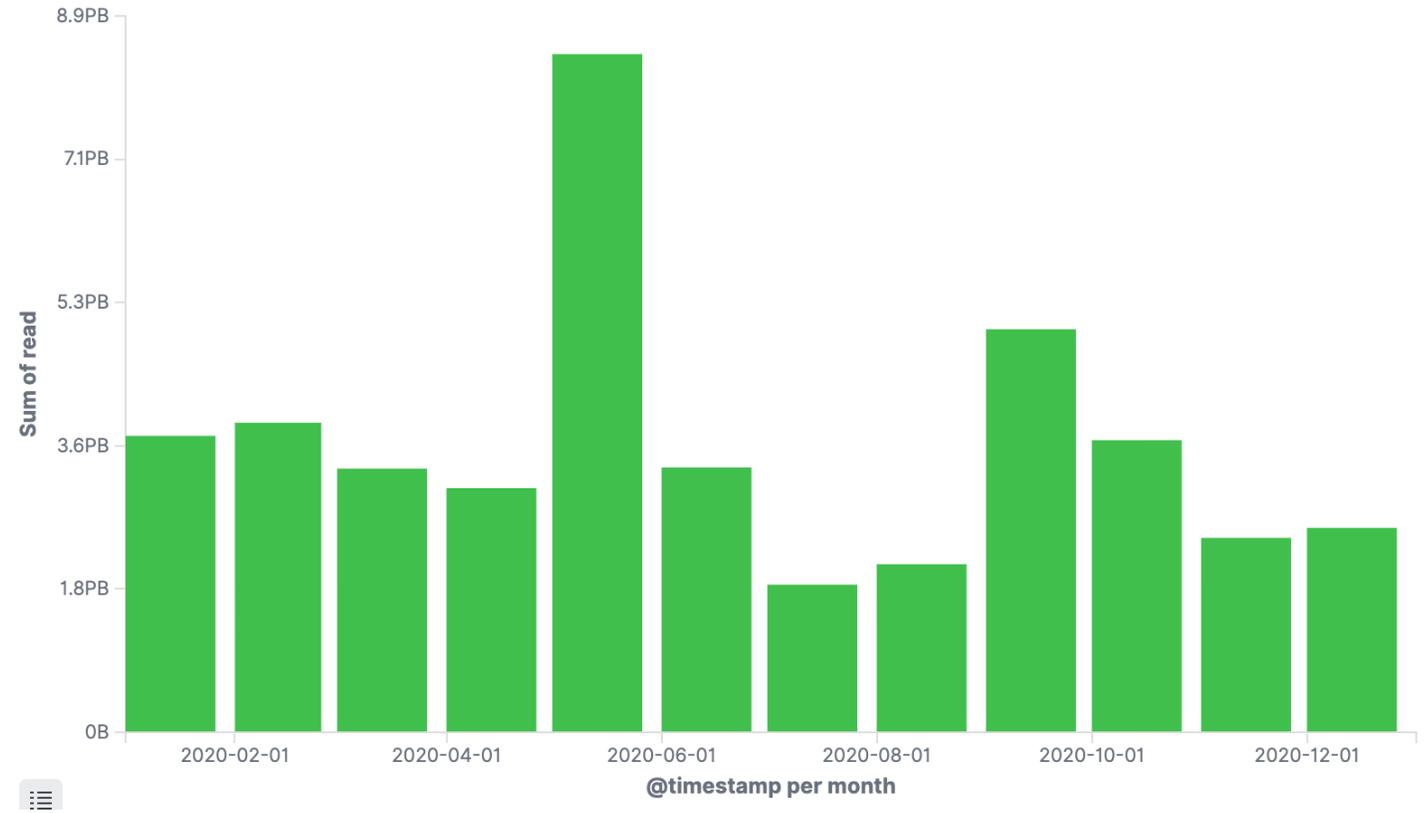
- IRIS-HEP requested validation of cache transfer accounting
- Completed correctness validation:
<https://doi.org/10.5281/zenodo.3981359>
- Scale validation is in-progress.
- Found some issues with missing accounting that is corrected with cache settings
- Will use the conclusions from the scale validation to inform how we deploy accounting

StashCache – Usage

- 2020:
 - Working set: **212TB**
 - Data read: **22PB**
 - OSGConnect Users: **84**

- February 2021:
 - Working set: **5.8TB**
 - Data read: **1.5PB**
 - OSGConnect Users: **17**

StashCache – 2020 Data Read



StashCache – Cache Usage February

- Internet2 caches are the most used
- Strategically placed throughout U.S: KC, Chicago, Manhattan, Sunnyvale

Cache / Origin	Data Read
osg.kans.nrp.internet2.edu	391.5TB
osg.chic.nrp.internet2.edu	301.1TB
osg.newy32aoa.nrp.internet2.edu	151TB
stashcache.t2.ucsd.edu	135.3TB
fiona-r-uva.vlan7.uvalight.net	64.5TB
osg.sunn.nrp.internet2.edu	51.8TB
sc-cache.chtc.wisc.edu	45.2TB
ds-102-11-18.cr.cnaf.infn.it	34.7TB
stashcache.gravity.cf.ac.uk	29.7TB

StashCache – Client Usage February

- Client domain may not be 100% accurate, difficult to capture

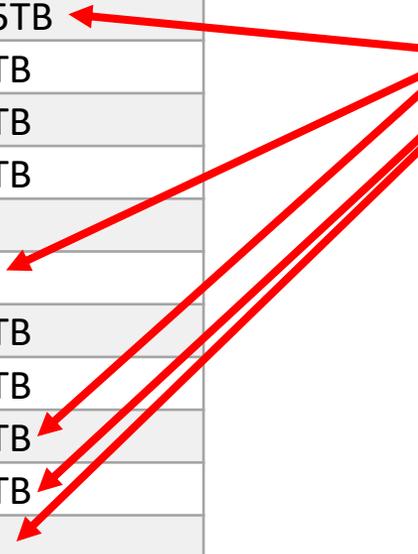
Client Domain	Bytes Read
caltech.edu	226.5TB
uconn.edu	54.8TB
unl.edu	48.4TB
illinois.edu	38.9TB
mwt2.org	36TB
infn.it	35TB
colorado.edu	28.3TB
aglt2.org	16.8TB
ac.uk	15.5TB
nikhef.nl	14.7TB
sdfarm.kr	14TB
amnh.org	11.2TB
iu.edu	10.4TB
gatech.edu	9.3TB

StashCache – Client Usage February

- Client domain may not be 100% accurate, difficult to capture

Client Domain	Bytes Read
caltech.edu	226.5TB
uconn.edu	54.8TB
unl.edu	48.4TB
illinois.edu	38.9TB
mwt2.org	36TB
infn.it	35TB
colorado.edu	28.3TB
aglt2.org	16.8TB
ac.uk	15.5TB
nikhef.nl	14.7TB
sdfarm.kr	14TB
amnh.org	11.2TB
iu.edu	10.4TB
gatech.edu	9.3TB

LIGO Usage

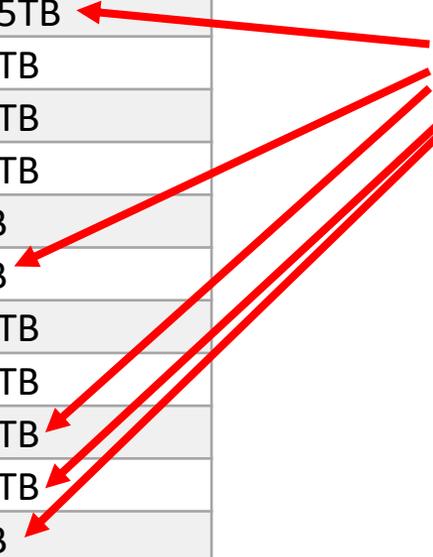


StashCache – Usage February

- LIGO is leading the way in exporting StashCache abroad!

Client Domain	Bytes Read
caltech.edu	226.5TB
uconn.edu	54.8TB
unl.edu	48.4TB
illinois.edu	38.9TB
mwt2.org	36TB
infn.it	35TB
colorado.edu	28.3TB
aglt2.org	16.8TB
ac.uk	15.5TB
nikhef.nl	14.7TB
sdfarm.kr	14TB
amnh.org	11.2TB
iu.edu	10.4TB
gatech.edu	9.3TB

LIGO Usage



StashCache – Usage February

- CC* Sites are showing up in clients
- StashCache means significant network utilization

Client Domain	Bytes Read
 caltech.edu	226.5TB
uconn.edu	54.8TB
unl.edu	48.4TB
illinois.edu	38.9TB
mwt2.org	36TB
infn.it	35TB
 colorado.edu	28.3TB
aglt2.org	16.8TB
ac.uk	15.5TB
nikhef.nl	14.7TB
sdfarm.kr	14TB
 amnh.org	11.2TB
iu.edu	10.4TB
 gatech.edu	9.3TB

Where to go from here?

- How do we make this easier for users?
 - Making transfers more transparent to the users?
- Continue to work with software providers to harden software and infrastructure
- Use the validation to find the holes in our accounting